An examination of the observed placebo effect associated with the treatment of low back pain – a systematic review

Aaron A Puhl MSc*, Christine J Reinhart PhD*, Elisabeth R Rok BSc*, H Stephen Injeyan PhD DC

L’examen de l’effet placebo observé associé au traitement des douleurs lombaires : une analyse systématique

OBJECTIF: Déterminer si les effets non spécifiques qui se produisent après l’utilisation d’interventions placebo pour traiter des douleurs lombaires non spécifiques (DLNS) sont suffisants pour être considérés comme significatifs sur le plan clinique.

MÉTHODOLOGIE : Les chercheurs ont effectué des recherches systématiques dans les bases de données électroniques afin de trouver des essais aléatoires contrôlés contre placebo sur des interventions de DLNS qui faisaient appel à des échographies placebo, au laser placebo et à des médicaments placebo comme contrôles. La sélection des études s’est effectuée au moyen d’une évaluation indépendante de l’admissibilité scientifique par trois réviseurs, et les décisions définitives d’inclusion étaient consensuelles.

RÉSULTATS : Aucune des études faisant appel à des échographies placebo comme contrôles dans le traitement des DLNS n’était acceptable à l’inclusion. Des résultats ont été incluses dans la présente évaluation de l’effet placebo, soit huit études qui respectaient les critères d’inclusion rigoureux de meilleures données probantes (trois faisant appel à un laser placebo et cinq, à des médicaments placebo) et quatre études sur des médicaments placebo qui « avaient été tout juste exclues » des critères d’inclusion de meilleures données probantes. Même si les données probantes tirées d’études faisant appel au laser placebo n’étaient pas concluantes, la présente analyse a découvert un changement significatif sur le plan clinique des indices de DLNS après l’utilisation de médicaments placebo par voie orale.

CONCLUSIONS : La présente analyse des meilleures données probantes a permis de découvrir un changement significatif des indices de douleur sur le plan clinique après l’utilisation de médicaments placebo par voie orale pour traiter des DLNS non spécifiques. Cette observation indique que d’autres recherches cliniques s’imposent pour déterminer quels sous-groupes de patients pourraient profiter le plus de ce traitement et pour distinguer le véritable apport de l’effet placebo de ses effets non spécifiques.

Key Words: Low back pain; Placebo; Systematic review; Therapy; Treatment

The placebo effect refers to the positive clinical outcomes caused by a treatment that is objectively without specific activity for the condition being treated (1). Placebos have long been reported to improve subjective outcomes in up to 35% of patients with a wide range of clinical conditions (2). Given that improvements have been reported to occur in such a sizable number of patients, it has been suggested that placebo treatments are an important tool that can be used by the medical community to complement regular therapies. However, in clinical medicine, the use of placebo treatments remains ethically controversial (3).

The use of placebo as a viable clinical intervention has been questioned by the findings of a recent meta-analysis by Hrobjartsson and Gøtzsche (4), including their use of meta-analysis to attempt to examine a complex phenomenon such as placebo. It has been suggested that meta-analyses are too mechanistic and are driven more by concerns of reproducibility rather than contributing to an understanding of the phenomenon of interest (5). More specifically, meta-analyses have the potential to dilute therapeutic outcomes, even possibly deem them insignificant, when heterogeneous therapies, diseases and outcomes are compiled in a single analysis (6). Quantifying the clinical efficacy of placebo is further complicated if one considers the suggestion by Moerman and Jonas (1) that the placebo effect is a “meaning response”, in which the placebo treatment is only effective if it has meaning to the patient. Thus, it is possible that subgroups within a patient population can respond differently to the same placebo intervention, dependent on whether they perceive it to be of

*These authors contributed equally to this work

Canadian Memorial Chiropractic College, Toronto, Ontario

Correspondence: Dr H Stephen Injeyan, Canadian Memorial Chiropractic College, 6100 Leslie Street, Toronto, Ontario M2H 3J1.

Telephone 416-482-2340 ext 172, fax 416-488-0470, e-mail sinjeyan@cmcc.ca

Pain Res Manage Vol 16 No 1 January/February 2011 ©2011 Pulsus Group Inc. All rights reserved 45
Puhl et al

great or little therapeutic value. Finally, discriminating the contribu-
tion of the true placebo effect from nonspecific effects such as natural
recovery and regression to the mean requires studies comparing pla-
cebo interventions with no-treatment controls, a methodology that
raises ethical concerns. Despite the controversy surrounding the clin-
ic efficacy of the placebo effect and the difficulty in studying this
phenomenon, it has been shown that approximately one-half of
internists and rheumatologists are prescribing placebo treatments
of some form on a regular basis, and that most physicians believe the
practice to be ethically permissible (7).

Low back pain (LBP) is a common condition, with a lifetime
prevalence as high as 84% (8) – second only to upper respiratory prob-
lems as a symptom-related reason for visits to a physician in the
United States (9). LBP can be defined as acute, subacute or chronic,
and the precise duration of each is heterogeneously defined through-
out the literature. It has often been stated that the natural history of
acute back pain has a favourable resolution, in part because early stud-
ies such as that by Coste et al (10) reported that 90% of LBP patients
seen within three days of onset recovered within two weeks. However,
investigations since then have suggested that it may be more accurate
to estimate that 25% to 30% of patients recover in one week, and 60%
of patients recover in seven weeks (11,12). A very recent study by
Henschke et al (13) found that nearly one-third of patients did not
recover from the presenting episode within a year. In line with these
individual findings, a recent systematic review (14) also indicated a
much less favourable prognosis than has commonly been reported. For
patients whose pain continues on to chronicity, back pain is rarely self-
limiting; fewer than 10% of patients with long-lasting back pain at
baseline reported no pain five years later (15).

Clinicians treating LBP may choose from a wide range of options
including exercise therapy, medications, manipulation, acupuncture,
electrical modalities and cognitive-behavioural interventions.
Currently, it appears that the ideal treatment for LBP is a multi-
disciplinary intervention with a stepwise approach (16) and, indeed,
studies examining the effectiveness of these treatments are becoming
numerous. In contrast, placebo treatments, although used on occasion
in clinical practice (7), have not been thoroughly investigated with
studies designed to specifically assess the placebo effect in the treat-
ment of LBP. This is particularly surprising when considering the
known contribution of placebo to the general effectiveness of anal-
gesic pharmaceuticals (17,18).

In an attempt to better understand how nonspecific effects might
contribute to the treatment of nonspecific LBP and to determine
whether studies designed specifically to assess placebo treatment of
LBP are warranted, a best-evidence model was used in the present
systematic review of randomized, controlled clinical trials. The goal of
the review was threefold – to first determine whether one treatment
results in a more profound change after a sham intervention than
another; second, to determine whether any of those observed effects
are large enough to be considered clinically meaningful; and finally, to
assess whether studies designed specifically to assess the contribution
of the true placebo effect, by comparison with no-treatment control
groups, are warranted.

METHODS

Definitions of placebo and LBP

In clinical trials, placebos are generally control treatments that are
indistinguishable from the study treatments but that lack their specific
activity. The most obvious way to blind patients to their allocation is
to provide one group with a placebo intervention that is structurally
equivalent to the experimental intervention. The intention of the
present review was to examine multiple forms of sham treatment, with
the requirement that subjects be completely blinded to the fact that
their treatment was functionally inert. With this in mind, sham ultra-
sound, sham cold laser and sham oral drug therapy were the only treat-
ments agreed on for inclusion in the present review (see Discussion).
Therefore, the examined studies were randomized clinical control

trials investigating treatments for nonspecific LBP that was acute
(ie, pain present for less than 48 h), subacute (ie, pain present between
two weeks and 12 weeks) or chronic (ie, pain present for longer than
three months) using sham laser, ultrasound or drug therapy as a pla-
cebo control. Nonspecific LBP was defined as pain below the 12th
costal margin and above the inferior gluteal folds, without pain radiat-
ing below the knee for which specific etiologies such as infection,
tumour, osteoporosis, fracture, structural deformity, inflammatory dis-
order, radicular syndrome or cauda equina syndrome, and other re-
levent pathological entities had been excluded.

Search strategy

A systematic search was conducted from January 1975 to December
2008 in MEDLINE (via PubMed), CINAHL, AMED (alternative medi-
cine), Nursing & Allied Health Collection: Comprehensive, and Alt
HealthWatch using the strategy recommended by the Cochrane Back
Review Group (19). The search was limited to English-language or
English language-translated human studies, and was repeated with syno-
nyms of “placebo” and “low back pain” and “randomized clinical trial”
and “laser therapy” or “ultrasound therapy” or “drug therapy”. Further
searches were performed with synonyms of “detuned ultrasound”,
“inactive laser” or “inert drug”. Separate searches were conducted for
each of the three therapies of interest. Selected literature reviews and
bibliographies were also examined to identify additional studies.

Study selection

Three reviewers (AP, CR and ER) were each assigned one or two
database(s) in which to apply the search strategy and locate poten-
tially relevant studies. Titles and abstracts were examined initially to
identify randomized placebo-controlled trials that used sham laser,
ultrasound or drugs to evaluate the efficacy of an intervention for LBP,
in which outcomes had been reported in terms of pain or disability.
Studies obtained from independent literature searches meeting the
inclusion criteria were pooled, and full texts were obtained.

The selected studies were subjected to a process of critical review
similar to that of the Task Force on Neck Pain (20), in which stan-
dardized forms were used to focus reviewers on the presence or absence
of important methodological issues. The critical review forms prompt-
reviewers to focus on issues such as study design, study population,
study conduct, participation rates, follow-up rates and measurement
issues. These key quality measures were consistent with the CON-
Solidated Standards of Reporting Trials (CONSORT) statement
(21) for ensuring the quality of randomized controlled trials. All four
authors independently read each article, and three (AP, CR and ER)
critically reviewed them independently. The entire team then met to
determine the articles to be included in the present review. Studies
were not assigned a quantitative score for judging scientific accept-
ability. Instead, the team discussed the internal validity of each study
and used consensus to decide whether a study was scientifically admis-
sible (22). An inadmissible study was one in which the methodological
weaknesses were judged to potentially lead to bias, such that the
study’s validity was significantly threatened and its results could not be
accepted with reasonable confidence. Because this synthesis was con-
cerned with the outcomes of the group receiving a placebo inter-
vention, emphasis was placed on judging methodology that could bias this
group’s outcomes.

To be included in the present systematic review, studies were required
to meet the following criteria: complete description of the LBP being
treated to ensure it conformed to the study definition of LBP; appro-
priate description of inclusion/exclusion criteria; blinding of the patients,
the person administering the therapy and/or the person assessing outcomes;
minimization of selection bias; absence of co-interventions; clear justifi-
cation of sample size; clear explanation of any subject withdrawals; and
clear indication of whether intention-to-treat analysis was performed
(eg, final analysis was based on outcomes from every subject who pro-
data at baseline and at least one postrandomization observation time).
**TABLE 1**

<table>
<thead>
<tr>
<th>Reason</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inclusion/exclusion criteria for low back pain examined outside the scope of this synthesis</td>
<td>(46-56)</td>
</tr>
<tr>
<td>Multiple complaints examined (eg, neck and back)</td>
<td>(49,57,58)</td>
</tr>
<tr>
<td>Multiple sham treatments or concurrent active intervention administered to placebo group</td>
<td>(59,60)</td>
</tr>
<tr>
<td>Insufficient description of inclusion/exclusion criteria</td>
<td>(57,58,61-66)</td>
</tr>
<tr>
<td>Possible confounding treatments allowed to continue or their exclusion not described (eg, chiropractic, acupuncture or massage therapy)</td>
<td>(32,35,48,62,64,67-70)</td>
</tr>
<tr>
<td>Lack of or inadequately described blinding</td>
<td>(64,68,71)</td>
</tr>
<tr>
<td>Weakness in blinding methodology (run-in phase)</td>
<td>(56,66,67,69,70)</td>
</tr>
<tr>
<td>Compliance or reasons for noncompliance inadequately described</td>
<td>(53,57,58,72)</td>
</tr>
<tr>
<td>Unclear if final outcomes based on intention-to-treat analysis</td>
<td>(67)</td>
</tr>
<tr>
<td>Comorbidities not excluded</td>
<td>(61,73)</td>
</tr>
<tr>
<td>Small sample size</td>
<td>(58,59,68)</td>
</tr>
<tr>
<td>Other</td>
<td>(51,52,64,72,73)</td>
</tr>
</tbody>
</table>

*Note: Most studies were found to have several weaknesses and, as such, are included in the table more than once*

---

Data synthesis

Data for the best-evidence synthesis consisted of studies that fulfilled the selection criteria and were, thus, judged to have adequate internal validity. The numerical data dealing with pain, disability or overall efficacy from the studies judged to be scientifically admissible were abstracted into tables by a single author (AP). For studies in which data were only provided in graphic format, an overlying grid was used to generate specific values. A meta-analysis of all studies or subgroups within the selected studies was not performed due to clinical and experimental heterogeneity. Study outcomes were individually examined for clinical relevance, in which a 30% or greater change from baseline was considered clinically meaningful, as previously established by consensus on clinical interpretation (23). In addition, studies were qualitatively analyzed according to the consistency of their evidence.

**RESULTS**

Literature search and quality appraisal

The initial search identified 658 titles/abstracts – 31 using sham ultrasound, 39 using sham laser and 588 using sham medication (Figure 1). Following the preliminary screening of these titles/abstracts for eligible studies, 40 nonduplicate studies were selected for further examination – two studies using sham ultrasound as a placebo, seven studies using sham laser as a placebo and 31 studies using sham oral medication as a placebo. Of the 40 studies selected for critical review, only eight were accepted as best evidence – three with sham laser placebo (24-26) and five with sham oral medication placebo (27-31). Although many of the remaining 32 studies contained multiple methodological weaknesses that resulted in their exclusion (Table 1), four studies using sham oral medication as a placebo were judged to have only one common methodological weakness (ie, failed to eliminate all possible confounding treatments). Because these studies were otherwise judged to have sufficient scientific validity, they were included in the final analysis as studies that ‘just missed’ being included as best evidence (32-35). No studies that used sham ultrasound for a placebo control in the treatment of LBP were acceptable for inclusion in the present review.

Qualitative analysis of best evidence

The characteristics and outcomes of the eight selected studies (24-31) are summarized in Tables 2 and 3. These studies primarily used subjective continuous outcome measures for both pain (eg, visual analogue scale [VAS]), McGill Pain Questionnaire and Modified Brief Pain Inventory) and disability (eg, Roland Morris Disability Questionnaire and Oswestry Disability Questionnaire). In addition, several studies used discrete outcome measures, such as patient and/or physician global assessment, including four- or five-point scales with word descriptors (eg, 0 = no relief, 1 = a little relief, 2 = some relief, 3 = a lot of relief and 4 = complete relief). The conclusions made in the present review were based primarily on the data from established continuous measures. For comparison purposes, discrete measures have also been presented in the study summary tables where applicable.

**Sham laser**

All three clinical trials using sham laser as a placebo intervention included subjects whose LBP could be considered chronic. A summary of the results of these studies is given in Table 2. Two studies (25,26) found no improvement of pain or disability from baseline after four weeks of sham laser intervention. However, the use of overall averages for reporting data in these studies may have concealed any minor benefits incurred by a small subgroup of their subject populations. One study (24) found a clinically meaningful improvement of pain in 69.7% of subjects in the placebo group following two weeks of sham laser intervention. Unfortunately, failure to consider data from 25% of patients who withdrew early from this study’s placebo group (due to lack of efficacy) certainly exaggerated the number of positive outcomes. Nevertheless, inclusion of the 25% of subjects who withdrew resulted in 55% of subjects experiencing a clinically relevant benefit from sham laser therapy.

**Sham medication**

Five studies using sham medication as a placebo intervention were judged to have sufficient scientific validity to be accepted as best evidence (Table 3) – two trials in acute LBP (27,28), one trial in subacute LBP (31) and two trials in chronic LBP (29,30). Both of the studies that included subjects with acute LBP (27,28) found a clinically meaningful improvement in average pain scores following three days of treatment. Only one of these studies examined disability (28), and although the authors found no significant improvement after three days, a meaningful improvement in average disability scores was observed after eight days. The single study that included subjects with subacute LBP (31) found clinically significant changes in average pain and disability scores after 30 days of treatment. However, at seven days of treatment, improvements in disability were not clinically significant, nor were pain outcomes reported at this time. Two studies included subjects with chronic LBP (29,30). Coats et al (30) found significant improvements in average pain scores after two and four weeks of treatment. Moreover, 41% of subjects in this study reported a greater than 50% improvement in pain after only one treatment.
### TABLE 2
Characteristics and outcomes of best-evidence trials that used a sham laser placebo

<table>
<thead>
<tr>
<th>Author, year (reference)</th>
<th>Symptom duration at baseline, inclusion criteria (mean)</th>
<th>Placebo group, n (completed)</th>
<th>Reasons for noncompliance</th>
<th>Age, years, mean ± SD</th>
<th>Characteristics of placebo treatment</th>
<th>Outcome measurement</th>
<th>Pain outcomes</th>
<th>Disability outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preyde, 2000 (26)</td>
<td>&gt;1 week, &lt;8 months (13.3 weeks)</td>
<td>27 (26)</td>
<td>NR</td>
<td>41.9±16.6</td>
<td>Sham laser treatment 6 times over 1 month</td>
<td>0 (baseline), 4 weeks and 8 weeks (follow up)</td>
<td>MPQ PRI score*: 2.0, 1.7 and 1.8; MPQ PRI score*: 11.1, 8.3 and 7.7</td>
<td>Roland Morris Disability Questionnaire score*: 7.2, 8.9 and 6.5; flexion ROM*: 5.5 cm, 6.0 cm and 5.5 cm</td>
</tr>
<tr>
<td>Basford et al, 1999 (25)</td>
<td>&gt;30 days (12.8 months)</td>
<td>29 (29)</td>
<td>NA</td>
<td>48.2±NR</td>
<td>Sham laser treatment 3 times per week for 4 weeks</td>
<td>0 (baseline), 2 weeks and 8 weeks (follow up)</td>
<td>VAS score* for maximum pain on palpation: 27.1 mm, 34.1 mm, 25.7 mm and 21.9 mm; VAS score* for maximum pain in previous 24 h: 37.4 mm, 38.6 mm, 32.8 mm and 35.1 mm</td>
<td>Oswestry Disability Index score*: 25, 22.9, 22.6 and 22.9; flexion ROM*: 14.2 cm, 14.4 cm, 14.0 cm and 14.2 cm</td>
</tr>
<tr>
<td>Soriano and Rios, 1998 (24)</td>
<td>&gt;3 months (NR)</td>
<td>42 (33)</td>
<td>64.3±NR</td>
<td>6.0 cm and 5.5 cm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Values represent mean scores. Bolded values represent a 30% or greater change from baseline. MPQ McGill Pain Questionnaire; NA Not applicable; NR Not reported; NSAID Nonsteroidal anti-inflammatory drug; PPI Present Pain Intensity; PRI Pain Rating Index; ROM Range of motion; VAS Visual analogue scale

### TABLE 3
Characteristics and outcomes of best-evidence trials that used a sham oral medication placebo

<table>
<thead>
<tr>
<th>Author, year (reference)</th>
<th>Symptom duration at baseline, inclusion criteria (mean)</th>
<th>Placebo group, n (completed)</th>
<th>Reasons for noncompliance</th>
<th>Age, years, mean ± SD</th>
<th>Characteristics of placebo treatment</th>
<th>Outcome measurement</th>
<th>Pain outcomes</th>
<th>Disability outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amlie et al, 1987 (27)</td>
<td>&lt;48 h and no previous episodes for 3 months (25.6 h)</td>
<td>142 (132)</td>
<td>10 withdrawn for protocol violations</td>
<td>38.1±113</td>
<td>Tablet per day for 7 days</td>
<td>0 (baseline), 3 days and 7 days</td>
<td>Improvement on VAS¹ for sitting: NA, 38% and 76%; lying: NA, 42% and 73%; standing: NA, 37% and 68%; walking: NA, 42% and 72%</td>
<td>Physicians’ opinion⁷ of functional improvement: NA, 40% and 79%</td>
</tr>
<tr>
<td>Dreiser et al, 2003 (28)</td>
<td>&lt;48 h and no previous episodes for 3 months (25.6 h)</td>
<td>126 (118)</td>
<td>8 withdrew due to adverse reactions</td>
<td>41.0±11.3</td>
<td>1 tablet per day for 7 days</td>
<td>0 (baseline), 3 days and 7 days</td>
<td>VAS scores¹: 71.4 mm, 46.0 mm and 33.3 mm; subjects rating pain relief ≥2 on patient global assessment⁸: NA, NR and 58%</td>
<td>RMDQ scores¹: 13, 9.3 and 7.3</td>
</tr>
<tr>
<td>Bannwarth et al, 2005 (31)</td>
<td>&gt;4 week but &lt;12 weeks (7.1 weeks)</td>
<td>80 (80)</td>
<td>NA</td>
<td>41.0±9.8</td>
<td>1 capsule per day for 30 days</td>
<td>0 (baseline), 3 days and 8 days and 30 days</td>
<td>VAS scores¹: 75 mm¹, 57 mm¹, 46 mm and 44 mm: Modified Brief Pain Inventory² interference composite scores: 39; 28, 23 and 22; subjects reporting &gt;50% pain relief: NA, 41%, NR and 54%</td>
<td>RMDQ scores¹: 11, 9.1 and 8.1</td>
</tr>
<tr>
<td>Coats et al, 2004 (30)</td>
<td>&gt;3 months requiring regular analgesic medication use (10.9 years)</td>
<td>145 (115)*</td>
<td>16 withdrew due to lack of efficacy; 14 due to protocol violation or other</td>
<td>48.7±12.6</td>
<td>1 tablet per day for 4 weeks</td>
<td>0 (baseline), 1 week, 2 weeks and 4 weeks; or time of withdrawal*</td>
<td>VAS scores¹: 75 mm¹, 57 mm¹, 46 mm and 44 mm: Modified Brief Pain Inventory² interference composite scores: 39; 28, 23 and 22; subjects reporting &gt;50% pain relief: NA, 41%, NR and 54%</td>
<td>RMDQ scores¹: 11, 9.1 and 8.1</td>
</tr>
<tr>
<td>Ruoff et al, 2003 (29)</td>
<td>&gt;3 months requiring daily analgesic medication (NR)</td>
<td>157 (74)*</td>
<td>59 withdrew due to lack of efficacy; 24 due to protocol violation or other</td>
<td>54.1±12.0</td>
<td>4–8 tablets per day for up to 91 days</td>
<td>0 (baseline)†, 1 week, 2 weeks and 4 weeks; or time of withdrawal*</td>
<td>VAS scores¹: 75 mm¹, 57 mm¹, 46 mm and 44 mm: Modified Brief Pain Inventory² interference composite scores: 39; 28, 23 and 22; subjects reporting &gt;50% pain relief: NA, 41%, NR and 54%</td>
<td>RMDQ scores¹: 11, 9.1 and 8.1</td>
</tr>
</tbody>
</table>

Bolded values represent a 30% or greater change from baseline. *Intention-to-treat analysis – outcomes were assessed at the final visit before withdrawal; †Baseline measurements were taken following a three-week medication washout phase; ²Values represent average scores; ³Pain relief was rated as 0 = no relief, 1 = a little relief, 2 = some relief, 3 = a lot of relief and 4 = complete relief; ⁴Efficacy of medication was rated as 0 = very poor, 1 = poor, 2 = good, 3 = very good and 4 = excellent. NA Not applicable; NR Not reported; RMDQ Roland Morris Disability Questionnaire; VAS Visual analogue scale

48 Pain Res Manage Vol 16 No 1 January/February 2011
week of treatment, which then increased to 54% of subjects reporting this level of improvement by the end of four weeks. Unlike pain, average disability scores only showed significant improvements after four weeks of treatment. The second study that examined chronic LBP (29) did not find a significant improvement in pain based on pooled subject averages. However, 39.5% of the subjects did report a clinically meaningful improvement in their individual pain score at the conclusion of 91 days of treatment. Furthermore, 32.5% of subjects individually reporting a clinically meaningful improvement in their pain scores showed a greater than 50% improvement. Average disability scores reported in this study did not indicate any clinically significant improvements (29).

Qualitative analysis of ‘just missed’ studies
All four studies included in the present review as having ‘just missed’ being included as best-evidence trials used sham oral medication as a placebo intervention for subjects with chronic LBP (32-35). A summary of the results of these studies is given in Table 4. One study found significant improvements in subjects’ average pain scores after two weeks of treatment (33) and two studies found significant improvements in subjects’ average pain scores after four weeks of treatment (32,34). Only one study (35) did not indicate a significant improvement in pain, based on pooled subject averages. However, 20.7% of subjects did report a greater than 30% improvement at the conclusion of 91 days of treatment. Two of the four ‘just missed’ studies (32,33) found significant improvements in average disability scores after four weeks.

| Author, year (reference) | 0 (baseline) | 1 week, 2 weeks, 4 weeks and 8 weeks (or time of withdrawal*) |
|--------------------------|---------------------------------|
| Peloso et al, 2004 (35)  | VAS score of how much pain in the previous week: | 51.9 mm |
| 12 weeks (or time of withdrawal*) | RMDQ score: 13.9†, 11.1, 10.6, 9.79, 9.9 and 9.79 |
| 0 (baseline) and 91 days (or time of withdrawal*) | RMDQ score: 15.0† and 13.7 |

DISCUSSION
The aim of the present review was threefold: first, to determine whether one treatment resulted in a more profound change in LBP symptoms after a sham intervention than another; second, to determine whether any of those observed effects were large enough to be considered clinically meaningful; and finally, to assess whether studies designed specifically to assess the contribution of the true placebo effect were warranted. While the initial intention of the present review was to examine multiple, prespecified forms of LBP treatment, the range of treatments qualifying for review was quickly limited because many forms of commonly used treatment (eg, chiropractic and acupuncture) have yet to be established and/or make sufficient use of an appropriate structurally equivalent and inert placebo (36,37). Furthermore, various electrotherapies were excluded because patients can experience a physical sensation during the application of most electrotherapeutic modalities. The difficulty in blinding subjects to placebo electrotherapy has been exemplified in a study by Deyo et al (38), in which multiple design features such as visually identical units, naive subjects, avoidance of a crossover design and use of identical visit frequency still only resulted in partial blinding success. Ultimately, ultrasound, laser and drug therapy were the only treatments found to meet the established criteria of the present review. Interestingly, however, very few LBP studies made use of ultrasound as a placebo, and none of those that did were admissible into our review. This was particularly surprising because sham or ‘subtherapeutic’ ultrasound has been suggested to be an effective placebo for use in LBP clinical trials (39).
Despite unanimous agreement among reviewers with regard to most methodological issues that resulted in the exclusion of studies, there was some debate that centred around two main methodological issues. The studies of greatest debate were methodologically sound for their respective inherent purposes, but each allowed the continuation of co-interventions that may have affected the very outcome measures that were the focus of the review. Four studies (32-35) allowed subjects to continue receiving manual therapies (ie, physical therapy, massage and chiropractic care) if their use had been stable for one to three months before entering the study. Although such treatments would not be expected to influence comparison studies after proper randomization, they do have the potential to bias the outcomes of each group independently. Despite this, these studies were included in the review as papers that ‘just missed’ being best evidence for several reasons. First, they all dealt with chronic LBP, which has an unfavourable long-term prognosis. More importantly, all co-interventions had been consistently used by subjects for one to three months before entering the study and, thus, baseline measurements could be expected to reflect the therapeutic value of these co-interventions.

The second methodological issue causing debate was the use of blinding, particularly in trials that used a run-in phase as part of their study. A run-in phase design means that all subjects were given an active treatment for a set time period and were then moved into a placebo-controlled double-blinded phase. Outcomes were measured as a regression of the initial improvements made during the run-in phase. This was considered to be a possible hindrance for successful blinding because of the potential for withdrawal symptoms to be experienced by the participants, which could ultimately make them aware of their placebo group designation. The potential for such complications was accentuated by the fact that all of the ‘run-in’ studies examined opioid medications, which are well known for creating dependency in patients (40). Consequently, these studies were rejected for the present review.

**Sham laser:** The present review found conflicting evidence with regard to the changes in pain and disability scores after the use of sham laser for the treatment of LBP or related disability. While one study found that a significant proportion of subjects had a greater than 30% improvement in pain (24), the remaining two studies showed no effect. Further extrapolation from these results was also hindered by the relatively small sample size used by studies deemed to be of high quality and accepted into the present review. Consequently, no firm conclusions could be drawn with regard to pain. For disability scores, none of the accepted studies found clinically meaningful improvements. Generally speaking, only a small number of studies examined the treatment of LBP in conjunction with using a sham laser placebo and, from the studies retrieved, most were judged to be inadequate in terms of scientific validity.

**Sham medication:** The present best-evidence review indicated that there is a clinically meaningful, observable change in pain scores following the use of sham oral medications for the treatment of non-specific LBP. Nine studies – all five of the best-evidence studies (27-31) and all four that ‘just missed’ this standard (32-35) – found a clinically significant improvement in average subjective, continuous pain scores after treatment when compared with baseline. Moreover, this improvement was reported in studies that examined acute, subacute and chronic LBP. In contrast, the effect of sham medications on disability was less favourable. Of the eight studies that examined disability, only five (28,30-33) found clinically meaningful improvements in subjective, continuous disability scores. Furthermore, because of the time course associated with this particular outcome measure, the possibility that the reductions in disability were simply a reflection of the natural history of the condition cannot be ruled out. Overall, the results suggest that the use of sham oral medications has a more pronounced effect on pain scores than disability scores. Unfortunately, based on the limited and heterogeneous nature of the data available in the present review, there is no indication of whether acute, subacute or chronic LBP would benefit most from the nonspecific effects associated with this sham intervention. In addition, due to the conflicting evidence found with regard to sham laser, a tangible comparison between sham laser and sham medication for the treatment of LBP is not possible; however, the results do suggest that the observed changes in pain associated with sham medication are more pronounced.

A noteworthy weakness of the present review was the inclusion of studies that lacked no-treatment control groups; thus, outcomes had to be based on comparisons with baseline measures. Indeed, patients who improved after receiving a placebo control may have improved as a result of the natural history of the disorder (41). Although it is irrefutable that the outcomes discussed herein represent an observed placebo response as opposed to a true placebo effect, it is difficult to differentiate the contribution of natural history (42). However, prognostic studies evaluating the natural progression of LBP are numerous and may be used as substitutes for the no-treatment control groups that are lacking in the included studies. In the present review, both studies concerning acute LBP (27,28) found an average improvement of 30% to 42% in VAS pain scores after only three days of treatment with sham medications, whereas natural history predicts that 25% to 30% of patients will improve in seven days (9). Even so, it remains difficult to conclude whether these improvements are due to natural progression of the condition.

More convincing evidence of the therapeutic relevance of sham medication may be provided from studies evaluating the treatment of chronic LBP. Four of the six chronic LBP studies reported in the present review found clinically meaningful improvements in average VAS pain scores after two to four weeks of treatment (30,32-34) and those that did not report such improvements in average pain scores reported meaningful improvements in a significant number of individual subjects (29,35). Moreover, the 30% improvement in average pain scores occurred in 40% to 54% of subjects who, at baseline, had reported daily analgesic use for up to 11 years (29,30). In fact, one study (30) reported a greater than 50% improvement in nearly one-half of the subjects after one week. Finally, several studies monitored subjects for as many as eight to 13 weeks (29,32,34,35), thus decreasing the likelihood that reported improvements are a convenient but temporary natural lull in symptoms.

Another common criticism of using baseline comparisons to evaluate clinical efficacy of placebo treatments is the inability to account for regression to the mean (4). In general, when observing repeated measurements in the same subject, relatively high observations of the outcome measure(s) are likely to be followed by less extreme ones at the next observation, with the latter being nearer to the subject’s true mean (43). All but two (27,34) of the oral medication studies included in the present review used baseline VAS pain scores of 40 mm or higher as an inclusion criterion. Thus, their subject populations are representative of an extreme subgroup and this has been suggested to increase the potential for regression to the mean (43). However, recent prognostic studies have indicated that higher LBP intensity at baseline is, in fact, associated with a poorer prognosis (13). Although it is impossible to quantify what the contribution of regression to the mean may have been, it is possible that using subjects with more intense LBP at baseline may have reduced the impact of this statistical phenomenon. It has been suggested that a treatment must have meaning to the patient before an effective placebo effect can be expected (1). It follows then that there is likely a subgroup of patients who would benefit more than others from the use of a particular placebo treatment. Indeed, when the average and individual changes in pain scores are compared, several studies presented in the present review do suggest that subgroups of patients are benefitting (29,30,35). Moreover, because other studies presented herein demonstrate only average changes in pain scores, it must be considered that, similarly, some of these patients are improving to a greater extent and others perhaps not at all. Regardless of the variability in improvements experienced by different patient subgroups, it has been shown that the effect of placebo analgesia can be greater when patients are told that a placebo treatment is actually a painkiller, compared with when they are...
told that they may receive either a painkiller or placebo (17,41). If there are no patient or clinician biases associated with the allocation of treatment, then the chances of the placebo group having a better outcome than the active treatment group are low (48). The observed placebo effect and low back pain

REFERENCES
